

GenCore version 4.5
Copyright (c) 1994 - 1999 GenCorp Inc.

OM protein - protein search, using sw model

Run on: January 22, 2000, 03:41:24, Search time 12.1 seconds
(without alignments)
64,172 Million cell updates/sec

Title: US-09-386-591-23

Perfect score: 134
Sequence: 1 RINGF11QM-FGFPRFIVFLQSLIS 26

Scoring table: BUCSUM62

Searched: 22229 seqs 22654216 residues

Database: SwissProt_231*

Word size: 0

Number of hits that pass the threshold: 82223

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	134	100.0	493	CE1P_HUMAN	Pl1597 homo sapien
2	134	100.0	493	CE1P_MACA	P47896 macaca fasc
3	131	97.6	289	CE1P_MESAU	P25941 mesocricetu
4	117	87.3	497	CE1P_PABIT	P22847 cryptotagu
5	52	23.9	217	YCAC_HAELN	C57121 haemophilu
6	50	37.3	311	YDAG_ECOLI	P76055 escherichia
7	48	35.8	348	AROH_EAELI	P06447 escherichia
8	46	34.7	377	QDS2_HFWSA	P05158 hemigrapsus
9	46	34.3	422	AP17_CABEL	P25522 caenorhabdi
10	46	34.3	422	AP17_CABEL	P25522 caenorhabdi
11	45.5	34.0	1246	YMW2_CABEL	P34504 caenorhabdi
12	45	33.5	774	P1F1_ACTAC	P26685 african swi
13	44.5	33.2	2290	P1G1_EMAC	P33354 encephalomy
14	44.5	33.2	2290	P1G1_EMAC	P33354 encephalomy
15	44.5	33.2	2290	P1G1_EMAC	P33354 encephalomy
16	44.5	33.2	901	P1G1_EMAC	P33354 encephalomy
17	44	32.9	429	AP50_CITET	P42540 mengo encop
18	44	32.9	472	QW01_APAAT	P44671 dihydropyri
19	44	32.9	472	QW01_APAAT	P44671 dihydropyri
20	44	32.9	472	QW01_APAAT	P44671 dihydropyri
21	43	32.5	340	QW01_APAAT	P44671 dihydropyri
22	43	32.5	340	QW01_APAAT	P44671 dihydropyri
23	43	32.5	340	QW01_APAAT	P44671 dihydropyri
24	43	32.5	340	QW01_APAAT	P44671 dihydropyri
25	43	32.5	340	QW01_APAAT	P44671 dihydropyri
26	43	32.5	340	QW01_APAAT	P44671 dihydropyri
27	43	32.5	340	QW01_APAAT	P44671 dihydropyri
28	43	32.5	340	QW01_APAAT	P44671 dihydropyri
29	43	32.5	340	QW01_APAAT	P44671 dihydropyri
30	43	32.5	340	QW01_APAAT	P44671 dihydropyri
31	43	32.5	340	QW01_APAAT	P44671 dihydropyri
32	43	32.5	340	QW01_APAAT	P44671 dihydropyri
33	43	32.5	340	QW01_APAAT	P44671 dihydropyri
34	43	32.5	340	QW01_APAAT	P44671 dihydropyri
35	43	32.5	340	QW01_APAAT	P44671 dihydropyri
36	43	32.5	340	QW01_APAAT	P44671 dihydropyri
37	43	32.5	340	QW01_APAAT	P44671 dihydropyri
38	43	32.5	340	QW01_APAAT	P44671 dihydropyri
39	43	32.5	340	QW01_APAAT	P44671 dihydropyri
40	43	32.5	340	QW01_APAAT	P44671 dihydropyri

ALIGNMENTS

Result No.	Score	Match	Length	ID	Description
1	134	100.0	493	CE1P_HUMAN	Pl1597 homo sapien
2	134	100.0	493	CE1P_MACA	P47896 macaca fasc
3	131	97.6	289	CE1P_MESAU	P25941 mesocricetu
4	117	87.3	497	CE1P_PABIT	P22847 cryptotagu
5	52	23.9	217	YCAC_HAELN	C57121 haemophilu
6	50	37.3	311	YDAG_ECOLI	P76055 escherichia
7	48	35.8	348	AROH_EAELI	P06447 escherichia
8	46	34.7	377	QDS2_HFWSA	P05158 hemigrapsus
9	46	34.3	422	AP17_CABEL	P25522 caenorhabdi
10	46	34.3	422	AP17_CABEL	P25522 caenorhabdi
11	45.5	34.0	1246	YMW2_CABEL	P34504 caenorhabdi
12	45	33.5	774	P1F1_ACTAC	P26685 african swi
13	44.5	33.2	2290	P1G1_EMAC	P33354 encephalomy
14	44.5	33.2	2290	P1G1_EMAC	P33354 encephalomy
15	44.5	33.2	2290	P1G1_EMAC	P33354 encephalomy
16	44.5	33.2	901	P1G1_EMAC	P33354 encephalomy
17	44	32.9	429	AP50_CITET	P42540 mengo encop
18	44	32.9	472	QW01_APAAT	P44671 dihydropyri
19	44	32.9	472	QW01_APAAT	P44671 dihydropyri
20	44	32.9	472	QW01_APAAT	P44671 dihydropyri
21	43	32.5	340	QW01_APAAT	P44671 dihydropyri
22	43	32.5	340	QW01_APAAT	P44671 dihydropyri
23	43	32.5	340	QW01_APAAT	P44671 dihydropyri
24	43	32.5	340	QW01_APAAT	P44671 dihydropyri
25	43	32.5	340	QW01_APAAT	P44671 dihydropyri
26	43	32.5	340	QW01_APAAT	P44671 dihydropyri
27	43	32.5	340	QW01_APAAT	P44671 dihydropyri
28	43	32.5	340	QW01_APAAT	P44671 dihydropyri
29	43	32.5	340	QW01_APAAT	P44671 dihydropyri
30	43	32.5	340	QW01_APAAT	P44671 dihydropyri
31	43	32.5	340	QW01_APAAT	P44671 dihydropyri
32	43	32.5	340	QW01_APAAT	P44671 dihydropyri
33	43	32.5	340	QW01_APAAT	P44671 dihydropyri
34	43	32.5	340	QW01_APAAT	P44671 dihydropyri
35	43	32.5	340	QW01_APAAT	P44671 dihydropyri
36	43	32.5	340	QW01_APAAT	P44671 dihydropyri
37	43	32.5	340	QW01_APAAT	P44671 dihydropyri
38	43	32.5	340	QW01_APAAT	P44671 dihydropyri
39	43	32.5	340	QW01_APAAT	P44671 dihydropyri
40	43	32.5	340	QW01_APAAT	P44671 dihydropyri

Thu Jan 27 11:30:43 2000

DR EMBL: M32494: AAA51978.1: JOINED.
 DR EMBL: M32495: AAA51978.1: JOINED.
 DR EMBL: M32496: AAA51978.1: JOINED.
 DR EMBL: M32497: AAA51978.1: JOINED.
 DR EMBL: M32573: AAB59388.1: -.
 DR PIR: A25941: A25941
 DR MIM: 118470: -.
 DR PROSITE: PS00400: LBP_RPL_CETP: 1.
 DR PFAM: PF01273: LBP_RPL_CETP: 1.
 KW Lipid transport; Cholesterol metabolism; glycoprotein; Signal;
 KW Atherosclerosis; Disease mutation.
 FT SIGNAL 1 17 CHOLESTERYL ESTER TRANSFER PROTEIN.
 FT CHAIN 18 493 POTENTIAL.
 FT CARBOHYD 105 105 POTENTIAL.
 FT CARBOHYD 257 257 POTENTIAL.
 FT CARBOHYD 358 358 POTENTIAL.
 FT CARBOHYD 413 413 POTENTIAL.
 FT VARIANT 459 459 ZPTG-VAR_004172.
 FT CONFLICT 422 422 I -> V (IN REF. 2).
 FT CAMELICT 251 310 MISSING (IN REF. 3).
 FT SEQUENCE 493 AA: 54770 MW: 59496080 CRC32:
 Query Match 100.0% Score 134: DB 1: Length 493:
 Best Local Similarity 100.0%: Pred. No. 2.5e-14
 Matches 26: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 Q7 1 RDGFLLQMDGFGPEHLLVDFLOSL 26
 DB 458 RDGFLLQMDGFGPEHLLVDFLOSL 493
 RESULT 2
 ID CETP_MACEA STANDARD: PRT: 493 AA.
 AC P47896:
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE CHOLESTERYL ESTER TRANSFER PROTEIN DEFENSIN (LIPID TRANSFER PROTEIN
 DE DEFENSIN).
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopitheciinae;
 OC Macaca.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 9204355.
 PA PAPE M.E., REHBERG F.F., MAROTTA K.R., MELCHIOR G.W.:
 FT "Molecular cloning, sequence, and expression of cynomolgus monkey
 FT cholesterol ester transfer protein, a large correlation between
 FT hepatic cholesterol ester transfer protein mRNA levels and plasma
 FT high density lipoprotein levels.";
 FT Arterioscler. Thromb. 11:1759-1771(1991).
 CC -1- FUNCTION: CETP IS AN EXTREMELY HYDROPHOBIC PROTEIN INVOLVED IN THE
 CC TRANSFER OF INSOLUBLE CHOLESTERYL ESTERS IN THE REVERSE TRANSPORT
 CC OF CHOLESTEROL.
 CC -1- SUBCELLULAR LOCATION: EXTRA-CELLULAR.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- SIMILARITY: BELONGS TO THE RPL_111 CETP FAMILY
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/usa> or
 CC or send an email to license@ebi.ac.uk).
 DR EMBL: M86443: AAA36840.1: -.
 DR PROSITE: PS00400: LBP_RPL_CETP: 1.

DR PFAM: PF01273: LBP_RPL_CETP: 1.
 KW Lipid transport; Cholesterol metabolism; glycoprotein; Signal;
 KW BY SIMILARITY.
 FT CHAIN 18 493 CHOLESTERYL ESTER TRANSFER PROTEIN.
 FT CARBOHYD 105 105 POTENTIAL.
 FT CARBOHYD 257 257 POTENTIAL.
 FT CARBOHYD 358 358 POTENTIAL.
 FT CARBOHYD 413 413 POTENTIAL.
 FT SEQUENCE 493 AA: 54743 MW: 4907620 CRC32:
 Query Match 100.0% Score 134: DB 1: Length 493:
 Best Local Similarity 100.0%: Pred. No. 2.5e-14
 Matches 26: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 Q7 1 RDGFLLQMDGFGPEHLLVDFLOSL 26
 DB 458 RDGFLLQMDGFGPEHLLVDFLOSL 493
 RESULT 2
 ID CETP_MESAU STANDARD: PRT: 289 AA.
 AC P5914:
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE CHOLESTERYL ESTER TRANSFER PROTEIN (LIPID TRANSFER PROTEIN 1)
 DE (FRAGMENT).
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Cricetidae; Mesocricetini.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 91154277.
 PA JIANG X.-C., MOULIN P., GUINEL E., GILBERT L., VANDER LIND
 RA "Mammalian adipose tissue and muscle are major sources of lipid
 RA transfer protein mRNA.";
 RA J. Biol. Chem. 265:531-546(1990).
 CC -1- FUNCTION: CETP IS AN EXTREMELY HYDROPHOBIC PROTEIN INVOLVED IN THE
 CC TRANSFER OF INSOLUBLE CHOLESTERYL ESTERS IN THE REVERSE TRANSPORT
 CC OF CHOLESTEROL.
 CC -1- SUBCELLULAR LOCATION: EXTRA-CELLULAR.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- SIMILARITY: BELONGS TO THE RPL_111 CETP FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/usa> or
 CC or send an email to license@ebi.ac.uk).
 DR EMBL: M63492: AAA37066.1: -.
 DR PIR: A38700: A38700.
 DR PROSITE: PS00400: LBP_RPL_CETP: 1.
 DR PFAM: PF01273: LBP_RPL_CETP: 1.
 KW Lipid transport; Cholesterol metabolism; glycoprotein;
 KW CARBOHYD 53 53 POTENTIAL.
 FT SEQUENCE 289 AA: 32430 MW: 3474774 CRC32:
 Query Match 97.8% Score 141: DB 1: Length 289:
 Best Local Similarity 96.2%: Pred. No. 4e-13
 Matches 25: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 Q5 1 RDGFLLQMDGFGPEHLLVDFLOSL 26
 DR EMBL: M86443: AAA36840.1: -.
 DR PROSITE: PS00400: LBP_RPL_CETP: 1.

```

DB 244 DQGLLMDGPPHLLVPLVLOS 284
RESULT 4
CETP_RABIT STANDARD; PRT: 497 AA.
AC F06637.
DI 01-AUG-1991 (Rel. 19, Created)
DI 01-AUG-1991 (Rel. 19, Last sequence update)
DI 01-NOV-1997 (Rel. 35, Last annotation update)
DE CHOLESTERYL ESTER TRANSFER PROTEIN PRECURSOR (CETP) TRANSFER PROTEIN
DE 1) (FRAGMENT)
GN CETP.
OS Oryctolagus cuniculus (Rabbit).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
CC Eutheria; Lagomorpha; Leporidae; Cuniculidae.
RN 1)
PP SEQUENCE FROM N.A.
PX MEDLINE: 89215620.
RA NAGASHIMA M., MEEHAN J.W., LAWN R.M.
RT "Cloning and expression of the rabbit cDNA for cDNA-activated protein
RT transfer protein."
RL J. Lipid Res. 32:1642-1643(1991).
CC 1) FUNCTION: CETP IS AN LARGELY HYDROPHOBIC PROTEIN INVOLVED IN THE
CC TRANSFER OF INCULATED CHOLESTERYL ESTERS IN THE REVERSE TRANSFER
CC OF CHOLESTEROL.
CC 2) SUBCELLULAR LOCATION: EXTRACELLULAR.
CC 3) TISSUE SPECIFICITY: PLASMA; SYNTHESIZED MAINLY IN THE LIVER.
CC 4) SIMILARITY: BELONGS TO THE LIPID-TRANSFER PROTEIN FAMILY.
CC 5) THIS SWISS-PROT entry is copyright. It is published through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL database.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Users of the database are
CC entitled to require a license agreement (See http://www.ebi.ac.uk/seqdb/).
CC 6) send an email to license@ebi.ac.uk.
DB 245 M27486; AAA1199.1; 1.
DB PROSITE: PS00400; LBP_PPT_CETP; 1.
DB PFAM: PF01274; LBP_PPT_CETP; 1.
KW Lipid transfer, Cholesterol metabolism, Glycoprotein, Signal
ET NONLIER 1
ET SIGNAL <1 1
ET CHAIN 2 497 CHOLESTERYL ESTER TRANSFER PROTEIN.
ET CARBOHYD 89 89 POTENTIAL.
ET CARBOHYD 99 99 POTENTIAL.
ET CARBOHYD 111 111 POTENTIAL.
ET CARBOHYD 241 241 POTENTIAL.
ET CARBOHYD 403 403 POTENTIAL.
ET CARBOHYD 417 417 POTENTIAL.
ET CARBOHYD 427 427 POTENTIAL.
S. SEQUENCE 127 AA. 54113 MW. 5281300 DAL. CRC32:
Query Match 82.3% Score 117. DB 1. Length 497.
Best Local Similarity 92.3%. Field No. 1.
Matches 23. Conservative 1. Mismatches 1. Indels 0. Gaps 0.
DB 473 DQGLLMDGPPHLLVPLVLOS 497
RESULT 5
YDAG_HABIN STANDARD. PRT: 313 AA.
AC Q57184; Q55459.
DI 01-NOV-1997 (Rel. 35, Created)
DI 01-NOV-1997 (Rel. 35, Last sequence update)
DI 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL PROTEIN H11371.1.
GN H11371.1.
OS Haemophilus influenzae.
CC Bacteria; Proteobacteria; gamma subphylum; Enterobacteriales;
CC Escherichia.
RN 1)
PP SEQUENCE FROM N.A.
PX STRAIN=K12 / MG1655;
RX MEDLINE: 97426517.
RA BLATTNER F., FUNKPIT G., JIN BUCH C.A., JENNA N.T., BURLAND V.,
RA BILEY M., GILLARD-VIDES J., GLASNER E.D., KODRICK M., MAYHEW G.F.,
RA GREGG P., TANTS N.W., PIPPALEPP P.A., JEFFER H.A., JOSE E.D.,
RA MAU P., SHAO Y.
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1233-1238(1997).
RN 1)
S. SEQUENCE OF 1-40 FROM N.A.
PX STRAIN=K12 / W3110;

```


PREVISIONS. SEQUENCE FROM N.A.
 STRAIN-ATCC 59477;
 MALDEN T. F., CLARK V. L., KURAMILSU H. K.;
 "Revised sequence of the Porphyromonas gingivalis pft cysteine
 protease/hemagglutinin gene: homology with streptococcal pyrokinic
 exotoxin E/streptococcal pyrogenase";
 Infect. Immun. 63:238-247(1995).
 -|- FUNCTION: APPEARS TO BE SPECIFIC FOR ARGININE CONTAINING PEPTIDE
 BONDS. POSSESSES HEMAGGLUTININ ACTIVITY.
 -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C10.
 -|- CAUTION: IT IS UNCERTAIN WHETHER MET 1 OR MET 17 IS THE INITIATOR.
 THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement. (See <http://www.ebi.ac.uk/edm/edm/edm/>)
 or send an email to license@isb-sib.ch.
 EMBL: M83096; -; NOT_ANNOTATED_CDS.
 PFAM: PF01643; pep_dhase_c10_1
 Hydrolase, Thiol, Peptide; Signal, Hemagglutinin.
 SIGNAL: 1 27 POTENTIAL.
 PROPEP 28 7
 CHAIN 7 868
 ACT_SITE 194 184
 ACT_SITE 127 127 BY SIMILARITY.
 SEQUENCE 968 AA: 96444 MW: 704245AE CRC32:

RESULT	11
ID	YMW2.CAPEL
STANDARD:	PRT: 1946 AA
AC	P44504; P44505; P34506; p60907;
BT	1-PEB-1994 (Rol. 28, Cited)
CI	15-DEC-1999 (Rol. 39, Last sequence update)
DE	15-DEC-1999 (Rol. 39, Last annotation update)
DT	HYPOTHETICAL 140.6 KD PROTEIN K04H4.2 IN CHROMOSOME 111.
K04H4.2.	
GN	Cacothabditis eloquans.
OR	Eukaryota; Metazoa; Nematoidea; Secernentea; Rhabdittia; Rhabditida;
OC	Rhabdittina; Rhabditoidea; Rhabditidae; Peleodermes; Cacothabditis.
LN	11.
RP	SEQUENCE FROM N.A.
RC	STRAIN BRISTOL #2.
EX	MEDLINE: 41510718.
KA	WILSON R., AINSFORTH K., ANDREWS E., RAYNER J., BEVIS M.
RA	BONFIELD S., BURTON J., CORNELL M., CORSEY T., COOPER J., COULSON A.
EA	CRANLEY M., DEAR S., DE DUDDIN B., FAVELLO A., FRASER A.
JA	FELLON L., GARDNER A., GREER P., HAWKINS I., HILLIER L., HIER M.,
RA	JACKSON L., LECHNING J., LLOYD C., MORITMORE B., O'CALLAGHAN M.,
RA	LAFELLE P., LICHNING J., RIKEN L., ROOPRA A., SAUNDERS E., STADEN R.,
PA	PARSONS J., PERCY C., SMITH A., SMITH M., SONNHAMMER E., STADEN R.,
PA	SIMS M., SWALDEN P., THOMPSON J., THOMAS K., VANTHER M., VAUGHAN E.,
KA	SULTON J., THIERRY-MIEG J., THOMAS K., VANTHER M., VAUGHAN E.,
KA	WATERSYN K., WATSON A., WEINSTOCK I., WILKINSON-SPROAT J.,
PA	WHILDEN P.;
PT	"2.2 MB OF CONTIGUOUS NUCLEOTIDE SEQUENCE FROM CHROMOSOME III OF C.
RI	eloquans";
RI	Nature 368:32-38(1994).
PN	121
RP	REVIEWS.

[illegible]


```

DR PFAM: PF00680: RNA_dep_RNA_pol: 1;
DR PFAM: PF00910: RNA_helicase: 1;
KW Polyprotein, Coat protein, Core protein, Translase,
KW RNA directed RNA polymerase, H2Oase, Thiol protease, Myristate,
FT PROPEP 1 67 LEADER PEPTIDE.
FT CHAIN 68 137 COAT PROTEIN VP4 (RHO).
FT CHAIN 148 393 COAT PROTEIN VP2 (BETA).
FT CHAIN 394 624 COAT PROTEIN VP3 (GAMMA).
FT CHAIN 625 901 COAT PROTEIN VP1 (ALPHA).
FT CHAIN 1059 1958 CORE PROTEIN P2A (G).
FT CHAIN 1194 1519 CORE PROTEIN P2B (I).
FT CHAIN 1520 1607 CORE PROTEIN P2C (F).
FT CHAIN 1608 1627 CORE PROTEIN P3A.
FT CHAIN 1628 1647 OLIGOMER-LINKED PROTEIN VP5 (H).
FT CHAIN 1648 1832 PICOIRNAIN 3C (P22).
FT CHAIN 1833 2292 RNA DIRECTED RNA POLYMERASE P3B (E).
FT LIPID 68 58 MYRISTATE (BY SIMILARITY).
FT ACT_SITE 1786 1786 PROTEASE (POTENTIAL).
FT ACT_SITE 1604 1604 PROTEASE (POTENTIAL).
SQ SEQUENCE 2292 AA: 255495 MW: 28095528 CRC32:

Query Match: 33.2% Score 44.56 DB 1: Length 2292
Best Local Similarity 52.4% Pos: 93, Length 2292.
Matches 11: Conservative 3: Mismatches 6: Indels 1: Gaps 1:

QY 1 RQGFLLIQMD-FGFPEHLVLD 20
111 111 1 11 11 1
DB 28 RRGFYLLFYDEEWYPEELTD 48

RESULT: 15
P016 EMV0
AC P17594: STANDARD: PRI: 2292 AA.
DT 01-AUG-1990 (Ref. 15, Created)
DT 01-FEB-1996 (Ref. 31, Last sequence update)
DT 15-DEC-1998 (Ref. 37, Last annotation update)
DE GENOME POLYPROTEIN (CONTAINS: COAT PROTEINS VP1 TO VP4, CORE PROTEINS
DE P2A TO P2C, P3A, GENOME LINKED PROTEIN VP5, PICOIRNAIN 3C
DE (EC 3.4.22.28) (G1-GLASE 3C) (P22). RNA DIRECTED RNA POLYMERASE P3B
DE (EC 2.7.7.48)1.
OS Encephalomyocarditis virus (strain end d diabetogenic).
OC Viruses; ssRNA positive strand viruses, no DNA stage, Picornaviridae;
OC Cardiovirus.
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE: 83243189.
RA BAE Y.S., EUN H.M., YOON J.W.:
RT "Genomic differences between the diabetogenic and nondiabetogenic
RT variants of encephalomyocarditis virus."
RL VIRIDITY 170:282:297(1989).
CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC 5'-5' ENDS IN THE P-REGION. IT MAY BE A CYSTEINE PROTEASE.
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS.
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- ITEM SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD NATURAL PEPTIDES.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C1.
CC This swiss-prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isdb.ch/submit/submit.htm)
CC or send an email to license@isb.ch.
CC -----
DR PFAM: M22458: AAA3034.1: 1;
DR PIR: A31474: GNNTED.
DR BSSP: P12296: LMEC.
DR PFAM: PF00073: RVV: 3.
DR PFAM: PF00680: RNA_dep_RNA_pol: 1.

```

```

DR PFAM: PF00910: RNA_helicase: 1;
KW Polyprotein, Coat protein, Core protein, Translase,
KW RNA directed RNA polymerase, Hydrolase, Thiol protease, Myristate,
FT PROPEP 1 67 LEADER PEPTIDE.
FT CHAIN 68 137 COAT PROTEIN VP4 (RHO).
FT CHAIN 148 393 COAT PROTEIN VP2 (BETA).
FT CHAIN 394 624 COAT PROTEIN VP3 (GAMMA).
FT CHAIN 625 901 COAT PROTEIN VP1 (ALPHA).
FT CHAIN 1059 1958 CORE PROTEIN P2A (G).
FT CHAIN 1194 1519 CORE PROTEIN P2B (I).
FT CHAIN 1520 1607 CORE PROTEIN P2C (F).
FT CHAIN 1608 1627 CORE PROTEIN P3A.
FT CHAIN 1628 1832 OLIGOMER-LINKED PROTEIN VP5 (H).
FT CHAIN 1833 2292 RNA DIRECTED RNA POLYMERASE P3B (E).
FT LIPID 68 58 MYRISTATE (BY SIMILARITY).
FT ACT_SITE 1786 1786 PROTEASE (POTENTIAL).
FT ACT_SITE 1604 1604 PROTEASE (POTENTIAL).
SQ SEQUENCE 2292 AA: 255426 MW: 1693194 CRC32:

```

```

Query Match: 33.2% Score 44.56 DB 1: Length 2292
Best Local Similarity 52.4% Pos: 93, Length 2292.
Matches 11: Conservative 3: Mismatches 6: Indels 1: Gaps 1:

QY 1 RQGFLLIQMD-FGFPEHLVLD 20
111 111 1 11 11 1
DB 28 RRGFYLLFYDEEWYPEELTD 48

```

Search completed: January 22, 2000, 03:46:09
Job time: 264 sec

.